REMARKS

Claims 1-35 are pending.

The Examiner rejected claims 1-35 under 35 U.S.C. 103(a) as being unpatentable over Damashek in view of Haimowitz et al.

Regarding claims 1, 12, 23, and 34, the Examiner stated that Damashek teaches the steps of: modifying the string of characters using a predetermined set of heuristics; responsive to not finding an exact match, forming a plurality of sub-strings of characters from the string of characters; using an information retrieval technique on the sub-strings of characters to determine a known string of characters equivalent to the string of characters; and comparing the modified string with a known string of characters in order to locate a match. The Examiner stated that Damashek does not teach performing a character-by-character comparison of the strings, but stated that Haimowitz teaches such a step at col. 5, line 66 to col. 6, line 1.

Applicants respectfully traverse this rejection.

Claim 1 recites:

> "A computer-implemented method of text equivalencing from a string of characters comprising:
>   modifying the string of characters using a predetermined set of heuristics;
>   performing a character-by-character comparison of the modified string with a known string of characters in order to locate a match;
>   responsive to not finding an exact match, forming a plurality of sub-strings of characters from the string of characters; and
>   using an information retrieval technique on the sub-strings of characters to determine a known string of characters equivalent to the string of characters."

The claimed method thus provides a technique for text equivalencing wherein a string is modified according to heuristics, and then the modified string is compared, on a character-by-character basis, with a known string to locate a match. If no match is found, a plurality of sub-strings is formed, and an information retrieval technique is used on the sub-strings to determine a known string of characters equivalent to the string.

Neither Damashek nor Haimowitz, nor any combination thereof, discloses the invention as claimed.

Damashek describes a technique for identifying, retrieving, or sorting documents by language or topic by creating an n-gram array for each document, parsing an unidentified document or query into n-grams, assigning a weight to each n-gram, removing commonality, comparing the n unidentified document or query to each document, scoring according to similarity, and, based on the similarity score, identifying, retrieving, or sorting the unidentified document or query with respect to language or topic. The comparison performed by Damashek, between the unidentified document or query and each document, is accomplished by determining an n-gram distribution for each, and comparing these distributions with one another. No string comparison is performed. The goal of Damashek is to determine a language or topic for the unidentified document or query; nowhere in Damashek is there any attempt to determine a text equivalent for a document or query, or to perform any type of string matching or comparison.

Haimowitz describes a technique for matching new customer records to existing customer records in a large database. This matching is performed by applying a field matching function to generate a score indicating the strength of a match between items in the field. Specifically, a hash key is applied to identify possible matches. The cited portion of Haimowitz, at col. 5, line 66 to col. 6, line 1, states, "The hash key function is composed of the first letter of the NAME field, a two-character COUNTRY_CODE, and the first seven characters of the CITY field." Thus, this cited portion describes a particular technique for generating a hash key from various attributes of a customer record.

The Examiner does not state how this cited portion of Haimowitz is related to the claimed invention. Rather, the Examiner states that it would have been obvious to one of ordinary skill in the art to perform a character-by-character comparison of strings before forming n-gram decomposition of strings in response to no match, because this would expedite the process of finding a matched string.

Applicants respectfully point out that the claimed invention is the combination of the four recited steps to perform a particular method of text equivalencing, and that this inventive combination is neither taught nor made obvious by the combination of Damashek and Haimowitz. In fact, neither of the cited references discloses any technique of text equivalencing whatsoever; moreover, both references are explicitly directed toward solving problems that are wholly distinct from those addressed by the present invention. To clarify this point: the present invention finds equivalent strings; Damashek classifies documents according to language or topic;

and Haimowitz generates and learns rules for matching database records. These are three wholly separate and distinct problems; the methods and steps for solving any one of these is not applicable to the others.

Furthermore, neither of the references discloses forming substrings responsive to not finding an exact match, as claimed herein. The Examiner states that such an operation would be obvious because this would expedite the process of finding a matched string; the Examiner's statement appears to set forth an advantage of the claimed invention but does not serve to indicate where in the prior art the claimed step is disclosed or rendered obvious.

Moreover, even were these two references combined in the manner suggested by the Examiner, the resulting combination would still fail to teach or disclose the limitations claimed herein. Damashek is directed toward classifying documents according to language or topic, while Haimowitz is directed toward matching new customer records to existing customer records. These are two completely different types of problems in different domains, and neither reference contains any hint or suggestion of the problem solved by the other reference. There is no teaching, nor indeed any hint or suggestion, within Damashek or Haimowitz of any combination of the two references, nor is there any discussion of any common problem or issue that such a combination would solve. Damashek compares n-gram distributions in order to classify documents; in the context of such an operation, comparison of individual strings to locate a match is a non sequitur operation that accomplishes nothing. The presence or absence of individual string matches is of no importance in the context of

Damashek; rather Damashek is comparing an overall <u>frequency distribution</u> of n-grams. On the other hand, Haimowitz is directed to generating rules for matching customer records to one another; the n-gram construction and frequency distribution comparison of Damashek is not operably combinable or applicable to such a technique.

The Examiner does not state where, in Damashek's process, a proposed step of comparing strings via a character-by-character method would be motivated or would even make operational sense. Applicants respectfully submit that the Examiner's proposed combination of the two references is not suggested by either reference, but rather derives from hindsight reconstruction of a solution to a problem that is discussed and solved by Applicants' own disclosure.

Claims 12 and 34 are system claims reciting elements corresponding to the method steps of claim 1. Claim 23 is a computer-readable medium claim comprising computer-readable code for performing the steps recited in claim 1. The other rejected claims various depend from claims 1, 12, 23, and 34, and introduce additional limitations that distinguish the invention from the cited art. The above arguments therefore apply to claims 2-35.

Applicants therefore submit that the claimed invention is patentably distinct over the combination of Damashek and Haimowitz.

In view of the above discussion, Applicants respectfully request that the §103 rejection be withdrawn. Favorable action is solicited.

Should the Examiner wish to discuss the above remarks, or if the Examiner believes that further contact with Applicants' representative would help to advance the prosecution of this case to finality, the Examiner is invited to telephone the undersigned at the number given below.

Respectfully submitted,
Ted E. Dunning and
Bradley D. Kindig

Dated: _MAR 24/04_

By: _____
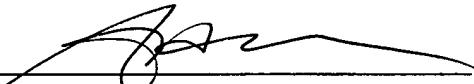
Amir H. Raubvogel,
Reg. No. 37,070
Fenwick & West LLP
801 California Street
Mountain View, CA 94041
Tel.: (650) 335-7276
Fax: (650) 938-5200